

Rate $\frac{1}{3}$ Index Coding: Forbidden and Feasible Configurations

V. Lalitha and Prasad Krishnan

SPCRC, International Institute of Information Technology, Hyderabad

Email: {lalitha.v, prasad.krishnan}@iiit.ac.in

Abstract—Linear index coding can be formulated as an interference alignment problem, in which precoding vectors of the minimum possible length are to be assigned to the messages in such a way that the precoding vector of a demand (at some receiver) is independent of the space of the interference (non side-information) precoding vectors. An index code has rate $\frac{1}{l}$ if the assigned vectors are of length l . In this paper, we introduce the notion of strictly rate $\frac{1}{L}$ message subsets which must necessarily be allocated precoding vectors from a strictly L -dimensional space ($L = 1, 2, 3$) in any rate $\frac{1}{3}$ code. We develop a general necessary condition for rate $\frac{1}{3}$ feasibility using intersections of strictly rate $\frac{1}{L}$ message subsets. We apply the necessary condition to show that the presence of certain interference configurations makes the index coding problem rate $\frac{1}{3}$ infeasible. We also obtain a class of index coding problems, containing certain interference configurations, which are rate $\frac{1}{3}$ feasible based on the idea of *contractions* of an index coding problem. Our necessary conditions for rate $\frac{1}{3}$ feasibility and the class of rate $\frac{1}{3}$ feasible problems obtained subsume all such known results for rate $\frac{1}{3}$ index coding.

I. INTRODUCTION

Index Coding, introduced in [1], considers the problem of efficiently broadcasting a set of messages available at a source, to a collection of receivers each of which already has a subset of the messages (called *side-information*) and demands certain other subset of messages. Based on the configuration of the messages available as side-information and the demand sets, the general index coding problem is classified into various types including unicast [2] (with disjoint demand sets), uniprior [3] (with disjoint side-information sets), and the most general *groupcast* index coding problems (arbitrary side-information and demand sets). The goal of index coding is to find optimal index codes, where optimality means that the number of channel uses is minimum. To determine the maximum rate (inversely related to minimum number of channel uses) of an index coding problem is NP-hard. The landmark paper [2] famously connected the scalar linear index coding problem to finding a quantity called *minrank* associated with the *side-information graph* related to the given single unicast index coding problem. The minrank problem is NP-hard too, but several approaches have been taken to address this problem, most popularly via graph theoretical ideas to bound the maximum rate (see, for example, [1], [2], [4]–[7]). The techniques used in these works to derive bounds on maximum rate naturally lead to constructions of (scalar and vector) linear index codes.

We focus on scalar linear codes in this paper. In [8], [9] the index coding problem has been reformulated in the interference alignment framework, where constructing a scalar

linear code was shown to be equivalent to assigning precoding vectors to the messages. The precoding vectors are assigned such that the space spanned by vectors assigned to interfering messages is linearly independent of the vector assigned to the demanded message. An index code is said to be a rate $\frac{1}{L}$ code (for some positive integer L) if the messages are scalars (from some field \mathbb{F}) and precoding vectors are L -length. In [8], a necessary and sufficient condition for index codes of rate $\frac{1}{2}$ was presented based on the structure of the interference (messages not available as side-information), which is modeled using a *conflict graph*. Consequently, a polynomial time algorithm to identify the feasibility of rate $\frac{1}{2}$ (or equivalently, minrank two) for a given index coding problem was given in [8].

Unlike the rate $\frac{1}{2}$ feasibility problem which has a polynomial time solution, the rate $\frac{1}{3}$ feasibility problem was shown to be NP-hard [10], when the finite field is fixed. A simple necessary condition for rate $\frac{1}{3}$ feasibility was given in [2]. A class of rate $\frac{1}{3}$ feasible index coding problems was shown in [8]. In [11], a stricter necessary condition was derived than what was previously known in [2]. In addition, a larger class of index coding problems which are rate $\frac{1}{3}$ feasible was presented, which include those given in [8]. Both the necessary and sufficient conditions obtained in [11] are based on the following two ideas (i) *conflict hypergraphs*, which preserve all the required information in the index coding problem (as opposed to conflict graph) (ii) *type-2 alignment sets*, which are special subsets of messages which must necessarily be assigned vectors from a two dimensional space.

In this work, we make further progress on the rate $\frac{1}{3}$ index coding problem. Our contributions are as follows.

- *Strictly rate $\frac{1}{L}$ feasible subsets*: We introduce the notion of strictly rate $\frac{1}{L}$ feasible subsets of messages of an index coding problem \mathbb{I} . A subset of messages is said to be strictly rate $\frac{1}{L}$ feasible if for any rate $\frac{1}{3}$ solution, there are always L linearly independent vectors amongst the precoding vectors assigned to the subset (Definition 8 in Section IV makes this precise).
- *Necessary condition based on strictly rate $\frac{1}{L}$ feasible subsets*: We show that a certain method of ‘stitching’ strictly rate $\frac{1}{2}$ feasible subsets together generates larger message subsets which should be strictly rate $\frac{1}{2}$ feasible. (Theorem 3, Section IV-B). Using this result, we recapture the main result of [11], while significantly generalizing it to include new configurations which must be strictly rate $\frac{1}{2}$. This, in turn, results in a general necessary condition for rate $\frac{1}{3}$ feasibility.
- *Contraction of an index coding problem*: We develop the

notion of a *contraction* of an index coding problem \mathbb{I} . A contraction of an index coding problem \mathbb{I} is another index coding problem \mathbb{I}' such that any solution for \mathbb{I}' gives a solution for \mathbb{I} (Section V gives a formal definition).

- *Sufficient condition based on contractions:* We give sufficient conditions on rate $\frac{1}{3}$ feasibility of \mathbb{I} based on the structure of type-2 alignments sets of a maximal contraction of \mathbb{I} (Theorem 4, Section V). These conditions result in a larger class of problems which are rate $\frac{1}{3}$ feasible.

Notations: Throughout the paper, we use the following notations. Let $[1 : m]$ denote $\{1, 2, \dots, m\}$. For a set of vectors A , $sp(A)$ denotes their span. For a vector space V , $dim(V)$ denotes its dimension. An arbitrary finite field is denoted by \mathbb{F} . A vector from the m -dimensional vector space \mathbb{F}^m is said to be picked *at random* if it is selected according to the uniform distribution on \mathbb{F}^m .

II. REVIEW OF INDEX CODING

Formally, the general index coding problem, called a *groupcast* index coding problem, consists of a broadcast channel which can carry symbols from \mathbb{F} , along with (i) A set of T receivers (ii) A source which has messages $\mathcal{W} = \{W_i, i \in [1 : n]\}$, each of which takes values from \mathbb{F} , (iii) For each receiver j , a set $D(j) \subseteq \mathcal{W}$ denoting the set of messages demanded by the receiver j , (iv) For each receiver j , a set $S(j) \subseteq \mathcal{W} \setminus D(j)$ denoting the set of side-information messages available at the j^{th} receiver. A *scalar linear index code* of symmetric rate $\frac{1}{L}$ (for some integer $L \geq 1$) for a given index coding problem consists of an encoding function $\mathbb{E} : \underbrace{\mathbb{F} \times \mathbb{F} \times \dots \times \mathbb{F}}_{n \text{ times}} \rightarrow \mathbb{F}^L$,

mapping the messages ($W_i \in \mathbb{F}$) to some L -length codeword which is broadcast through the channel, as well as decoding functions at the receivers which use the codeword and the available side-information symbols to decode the demands. The encoding function of a scalar linear index code can be expressed as $\mathbb{E}(W_1, W_2, \dots, W_n) = \sum_{i=1}^n V_i W_i$, where V_i is a L -length vector over \mathbb{F} called the precoding vector assigned to W_i . Finding a scalar linear index code of length L (i.e., with a feasible rate $1/L$) is equivalent to finding an assignment of these V_i s to the n messages such that the receivers can all decode their demanded messages.

Definition 1 (Interfering sets and messages, conflicts). *For some receiver j and for some message $W_k \in D(j)$, let $Interf_k(j) \triangleq \mathcal{W} \setminus (W_k \cup S(j))$ denote the set of messages (except W_k) not available at the receiver j . The sets $Interf_k(j), \forall k$ are called the interfering sets at receiver j . If receiver j does not demand message W_k , then we define $Interf_k(j) \triangleq \emptyset$. If a message set \mathcal{W}_i is not available at a receiver j demanding at least one message $W_k \notin \mathcal{W}_i$, then \mathcal{W}_i is said to interfere at receiver j , and \mathcal{W}_i and W_k are said to be in conflict.*

For a set of vertices $A \subseteq \mathcal{W}$, let $V_{\mathbb{E}}(A)$ denote the vector space spanned by the vectors assigned to the messages in A , under the specified encoding function \mathbb{E} . If $A = \emptyset$, we define $V_{\mathbb{E}}(A)$ as the zero vector. A message subset A is said to be L -dimensional (under the code \mathbb{E}) if $dim(V_{\mathbb{E}}(A)) = L$.

Definition 2 (Resolved conflicts). *For a given assignment of vectors to the messages (or equivalently, for a given encoding function \mathbb{E}), we say that all the conflicts are resolved if $V_k \notin V_{\mathbb{E}}(Interf_k(j)), \forall W_k \in D(j), \forall j$.*

It can easy to show that successful decoding at the receivers is possible if and only if all the conflicts are resolved [11]. We alternatively refer to a collection of messages by only their indices (for example, $\{W_1, W_2\}$ is referred to as $\{1, 2\}$).

III. PRIOR RESULTS FROM [11] ON RATE $\frac{1}{3}$ INDEX CODES

In [11], we developed a new framework for studying the rate $\frac{1}{3}$ feasibility of groupcast index coding problems. We recall the basic definitions and results from [11]. We build on these results in this paper.

Definition 3 (Conflict hypergraph). *The conflict hypergraph is an undirected hypergraph with vertex set \mathcal{W} (the set of messages), and its hyperedges defined as follows.*

- *For any receiver j demanding any message W_k , W_k and $Interf_k(j)$ are connected by a hyperedge (shown dotted in our figures), which is denoted by $\{W_k, Interf_k(j)\}$.*

It was shown in [11] that the conflict hypergraph sufficiently captures the essence of the index coding problem.

Definition 4 (Alignment graph and alignment sets - [8]). *The alignment graph is an undirected graph with vertex set \mathcal{W} and edges defined as follows. The vertices W_i and W_j are connected by an edge (called an alignment edge, shown in our figures by a solid edge) when the messages W_i and W_j are not available at a receiver demanding a message other than W_i and W_j . A connected component of the alignment graph is called an alignment set.*

Definition 5 (Restricted Index Coding problem). *Let \mathbb{I} denote an index coding problem with message set \mathcal{W} . For some $\mathcal{W}' \subseteq \mathcal{W}$, a \mathcal{W}' -restricted index coding problem is defined as the index coding problem $\mathbb{I}_{\mathcal{W}'}$ consisting of (i) The messages \mathcal{W}' , (ii) The subset $\mathcal{T}_{\mathcal{W}'}$ of the receivers of \mathbb{I} which demand messages in \mathcal{W}' , (iii) For each $j \in \mathcal{T}_{\mathcal{W}'}$ the demand sets $D_{\mathcal{W}'}(j)$ and the side-information sets $S_{\mathcal{W}'}(j)$ are restricted within \mathcal{W}' , i.e., $D_{\mathcal{W}'}(j) = D(j) \cap \mathcal{W}'$ and $S_{\mathcal{W}'}(j) = S(j) \cap \mathcal{W}'$.*

The alignment graph and the alignment sets of the restricted index coding problem $\mathbb{I}_{\mathcal{W}'}$ are called the \mathcal{W}' -restricted alignment graph and \mathcal{W}' -restricted alignment sets respectively. A \mathcal{W}' -restricted internal conflict is a conflict between any two messages within a restricted alignment set of \mathcal{W}' .

Definition 6 (Triangular Interfering Sets, Adjacent Triangular Interfering Sets). *A subset $\mathcal{W}'' \subset \mathcal{W}$ of size three is said to be a triangular interfering set if all the messages in \mathcal{W}'' interfere simultaneously at some receiver, and at least two of the messages in \mathcal{W}'' are in conflict. Two distinct triangular interfering sets \mathcal{W}_1 and \mathcal{W}_2 are said to be adjacent if $\mathcal{W}_1 \cap \mathcal{W}_2 = \{W_i, W_j\}$ such that W_i and W_j are in conflict.*

Definition 7 (Connected triangular interfering sets, Type-2 alignment sets). *Two triangular interfering sets \mathcal{W}_1 and \mathcal{W}_2*

are said to be connected if there exists a path (i.e., a sequence) of adjacent triangular interfering sets starting from \mathcal{W}_1 and ending at \mathcal{W}_2 . A type-2 alignment set is a maximal set of triangular interfering sets which are connected to each other.

The following theorem gives a necessary and sufficient condition for assigning vectors from a two dimensional space to the type-2 alignment sets.

Theorem 1. *Let \mathcal{W}' be a type-2 alignment set of the given index coding problem \mathbb{I} . If \mathbb{I} is rate $\frac{1}{3}$ feasible, then $\mathbb{I}_{\mathcal{W}'}$ must be rate $\frac{1}{2}$ feasible which holds if and only if there are no \mathcal{W}' -restricted internal conflicts.*

Theorem 2. *A rate $\frac{1}{2}$ infeasible index coding problem \mathbb{I} is rate $\frac{1}{3}$ feasible if every alignment set of \mathbb{I} satisfies either of the following conditions.*

- 1) It does not have both forks and cycles (a fork is a vertex connected by three or more edges).
- 2) It is a type-2 alignment set with no restricted internal conflicts.

IV. A GENERAL NECESSARY CONDITION FOR RATE $\frac{1}{3}$ FEASIBILITY

We begin with the formal definition of strictly rate $\frac{1}{L}$ feasible and infeasible subsets.

Definition 8 (Strictly rate $\frac{1}{L}$ feasible and infeasible subsets). *Let \mathbb{I} be an index coding problem with message set \mathcal{W} . A subset $\mathcal{W}' \subseteq \mathcal{W}$ is said to be strictly rate $\frac{1}{L}$ feasible (or simply, strictly rate $\frac{1}{L}$) if for any rate $\frac{1}{3}$ code given by an encoding function \mathbb{E} (if it exists), we have $\dim(\mathbb{V}_{\mathbb{E}}(\mathcal{W}')) = L$. A subset \mathcal{W}' is called strictly rate $\frac{1}{L}$ infeasible if $\dim(\mathbb{V}_{\mathbb{E}}(\mathcal{W}')) \neq L$ for any rate $\frac{1}{3}$ encoding function \mathbb{E} .*

Examples of strictly rate $\frac{1}{2}$ subsets include triangular interfering sets and type-2 sets. In the forthcoming part of this paper, we also construct other classes of strictly $\frac{1}{2}$ subsets as well as examples of strictly rate 1 and rate $\frac{1}{3}$ subsets. Some of such sets help us to further characterize index coding problems which are rate $\frac{1}{3}$ feasible based on the existence of some substructures, like the type-2 set conditions given by Theorem 1.

A. Two Interference Configurations and their strictly rate $\frac{1}{L}$ feasible (infeasible) subsets

Here, we will present two interference configurations (i) Successive triangular interference configuration (STIC) and (ii) Square pyramid interference configuration (SPIC). We will identify strictly rate $\frac{1}{L}$ feasible (and infeasible) subsets in these two configurations. We will also present an example based on two STIC subsets which is rate $\frac{1}{3}$ infeasible.

Definition 9 (Successive Triangular Interference Configuration (STIC)). *A collection of messages $\mathcal{W}_{STIC} = \{1, 2, 3, 4, 5, 6\}$ is said to be a successive triangular interference configuration (STIC) subset (Fig. 1(a)) if the following conditions hold.*

- The sets $\{1, 2, 3\}$, $\{2, 4, 5\}$, and $\{3, 5, 6\}$ interfere at some receivers demanding messages other than $\{1, 2, 3, 4, 5, 6\}$.

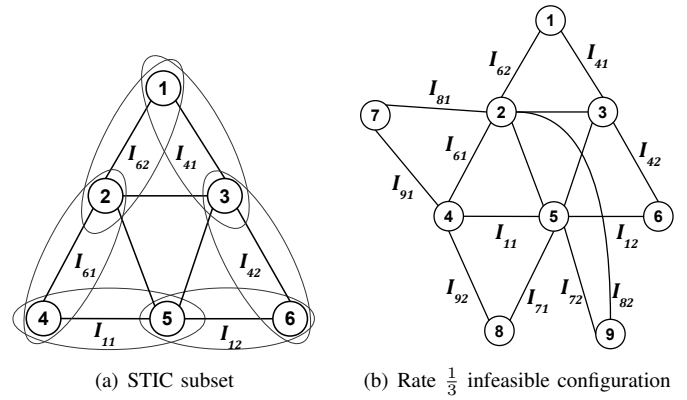


Fig. 1. Fig. 1(a) shows a STIC subset. Messages in the ellipses I_{ij} are interfering sets at some receiver demanding W_i . Fig. 1(b) shows a rate $\frac{1}{3}$ infeasible configuration based on STIC sets. The sets $\{1, 2, 3, 4, 5, 6\}$ and $\{7, 4, 2, 8, 5, 9\}$ are both STIC sets. The messages along the edges marked I_{ij} are interfering sets at a receiver demanding W_i .

- The sets $\{1, 2\}$ and $\{2, 4\}$ interfere at W_6 , $\{1, 3\}$ and $\{3, 6\}$ interfere at W_4 , and $\{4, 5\}$ and $\{5, 6\}$ interfere at W_1 (and there are no other conflicts amongst $\{1, 2, 3, 4, 5, 6\}$).

Lemma 1. *Let $\{1, 2, 3, 4, 5, 6\}$ be a STIC set as in Fig. 1(a). Then the following statements are true.*

- 1) The set $\{2, 3, 5\}$ is strictly rate $\frac{1}{2}$ infeasible.
- 2) The sets $\{1, 2, 3\}$, $\{2, 4, 5\}$, and $\{3, 5, 6\}$ are strictly rate $\frac{1}{2}$ feasible.

Proof: Proof of 1): Suppose $\{2, 3, 5\}$ is two dimensional for some rate $\frac{1}{3}$ index code given by \mathbb{E} . WLOG, let us assume that V_2 and V_3 are linearly independent, and thus $V_5 \in \text{sp}\{V_2, V_3\}$. Furthermore, $V_1 \in \text{sp}\{V_2, V_3\}$ as $\{1, 2, 3\}$ is interfering at some receiver.

Now, if $V_5 = \alpha V_2 + \beta V_3$ for some $\alpha, \beta \in \mathbb{F}$ both non-zero, then it must be that $V_4 = V_5$ (or a scalar multiple). Otherwise V_4 and V_5 will be independent, and hence we will have $sp(\{V_2, V_3\}) = sp(\{V_2, V_5\}) = sp(\{V_4, V_5\})$ and thus $V_1 \in sp(\{V_4, V_5\})$, which is not allowed. Similarly, we should have $V_6 = V_5$ too. This means that $V_4 \in sp(\{V_3, V_6\})$, which is not allowed. Hence we cannot have $V_5 = \alpha V_2 + \beta V_3$ with α, β both non-zero.

Thus, WLOG, let $V_5 = V_2$ (or equivalently, some constant multiple). Suppose V_6 (which is in $sp\{V_3, V_5\}$) as V_3 and V_5 are independent) and V_5 are independent. Then $V_1 \in sp\{V_5, V_6\}$ which is not allowed. Thus $V_6 = V_5 = V_2$. However this means that $V_6 \in sp\{V_2, V_1\}$ for any choice of V_1 , which is not allowed again. A similar contradiction arises if $V_5 = V_3$. This concludes proof of 1).

Proof of 2): Clearly as the sets $\{1, 2, 3\}$, $\{2, 4, 5\}$, and $\{3, 5, 6\}$ are all interfering sets, we must have that each of them is strictly rate $\frac{1}{3}$ infeasible. Suppose $\{1, 2, 3\}$ is one dimensional. Then $\{2, 3, 5\}$ must also be one dimensional as $\{2, 3, 5\}$ cannot be two dimensional by 1). But then, $V_1 = V_5$ which is not allowed. Thus $\{1, 2, 3\}$ must be two dimensional, and a similar argument applies for the other two sets. ■

The following example, based on STIC sets, gives a rate

$\frac{1}{3}$ infeasible problem based on an alignment structure constructed using two STIC sets, but does not have any triangular interfering sets (unlike all known previous examples). This also suggest the hardness of deciding the feasibility of rate $\frac{1}{3}$ index coding.

Example 1. Consider an index coding problem \mathbb{I} containing the alignment set structure shown in Fig. 1(b). There are two STIC sets $\{1, 2, 3, 4, 5, 6\}$ and $\{7, 4, 2, 8, 5, 9\}$ which share the common set $\{2, 4, 5\}$. Consider the index coding problems \mathbb{I}_1 (respectively, \mathbb{I}_2) restricted to the STIC set $\{1, 2, 3, 4, 5, 6\}$ (equivalently, $\{7, 4, 2, 8, 5, 9\}$) and all the messages at which these messages interfere. Clearly, by Lemma 1, $\{2, 4, 5\}$ is strictly rate $\frac{1}{2}$ feasible in \mathbb{I}_1 , while the same set is strictly rate $\frac{1}{2}$ infeasible in \mathbb{I}_2 . Note that any rate $\frac{1}{3}$ feasible code for \mathbb{I} is also feasible for \mathbb{I}_1 and \mathbb{I}_2 . This means $\{2, 4, 5\}$ is simultaneously two dimensional and also not so. This is meaningless, hence we have that \mathbb{I} is rate $\frac{1}{3}$ infeasible.

Remark 1. As a general rule, if there is a subset of messages which is strictly rate $\frac{1}{L}$ infeasible in a restricted index coding problem of the given problem, and strictly rate $\frac{1}{L}$ feasible in another restricted problem, then the given problem must clearly be rate $\frac{1}{3}$ infeasible.

We now define the SPIC sets, using which we obtain a new class of strict $\frac{1}{2}$ feasible sets.

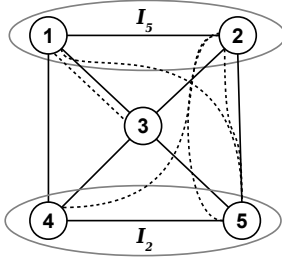


Fig. 2. A square pyramid interference configuration (SPIC) set: The receivers, at which the sets $\{1, 2, 3\}$, $\{1, 3, 4\}$, $\{3, 4, 5\}$ and $\{2, 3, 5\}$ are conflicting, are suppressed. Definition 10 has the complete details.

Definition 10 (Square Pyramid Interference Configuration (SPIC)). A collection of messages $\mathcal{W}_{SPIC} = \{1, 2, 3, 4, 5\}$ is said to be a square pyramid interference configuration (SPIC) set (see Fig. 2) if the following conditions hold.

- Message sets $\{1, 2, 3\}$, $\{1, 3, 4\}$, $\{3, 4, 5\}$ and $\{2, 3, 5\}$ interfere at some receivers which do not demand the messages $\{1, 2, 3, 4, 5\}$.
- Message set $\{1, 2\}$ (denoted by I_5) interferes at a receiver demanding W_5 , and $\{4, 5\}$ (denoted by I_2) interferes at a receiver demanding W_2 .
- Messages W_1 and W_3 are in conflict.
- No conflicts other than the above exist between the messages $\{1, 2, 3, 4, 5\}$.

We now show that any SPIC set has to be strictly rate $\frac{1}{2}$, and also give other properties which it must satisfy, in the following lemma.

Lemma 2. A SPIC set \mathcal{W}_{SPIC} as shown in Fig. 2 satisfies the following conditions.

- 1) The triangular interfering sets $\{1, 2, 3\}$, $\{1, 3, 4\}$, $\{2, 3, 5\}$ and the set $\{1, 2, 3, 4\}$ are strictly rate $\frac{1}{2}$.
- 2) \mathcal{W}_{SPIC} is strictly rate $\frac{1}{2}$.
- 3) The sets $\{1, 4\}$, and $\{2, 5\}$ are strictly rate $\frac{1}{2}$.
- 4) The sets $\{4, 5\}$ and $\{1, 2\}$ must be strictly rate 1.

Proof: Proof of 1): This follows by Theorem 1 as $\{1, 2, 3, 4\}$ must be a subset of a type-2 alignment set, and also because $\{2, 3, 5\}$ is a triangular interference set. Thus, we must have $V_{\mathbb{E}}(\{1, 2, 3, 4\}) = V_{\mathbb{E}}(\{1, 3\})$ for any rate $\frac{1}{3}$ feasible assignment \mathbb{E} .

Proof of 2): Let \mathbb{E} be the encoding function of some rate $\frac{1}{3}$ index code. Clearly $2 \leq \dim(V_{\mathbb{E}}(\mathcal{W}_{SPIC})) \leq 3$. Suppose that $\dim(V_{\mathbb{E}}(\mathcal{W}_{SPIC})) = 3$. By 1), we must therefore have that V_5 is independent of the space $V_{\mathbb{E}}(\{1, 2, 3, 4\})$. Clearly, V_2 and V_4 must be linearly independent, as W_2 and W_4 are in conflict. By 1), we have $V_3 \in \text{sp}(\{V_4, V_2\})$. It is easy to see that any such assignment for V_3 makes at least one of the sets $\{3, 4, 5\}$ and $\{2, 3, 5\}$ as three dimensional, which is not allowed. This proves 2).

Proof of 3) and 4): By 2), it must be that $\{1, 2, 5\}$ must also be assigned vectors from the space spanned by $\{V_1, V_3\}$. Now, if V_1 and V_2 are linearly independent, then the conflicts will not be resolved at W_5 . Thus, we must have $V_2 = V_1$ (or a constant multiple). By the similar arguments, we must have $V_4 = V_5 = \alpha V_1 + \beta V_3$ (for some $\alpha, \beta \in \mathbb{F}, \beta \neq 0$). This is the only assignment which is possible and it can be checked that 3) and 4) hold. ■

B. Necessary condition based on strictly rate $\frac{1}{2}$ feasible subsets

In the following theorem, we show that if the strictly rate $\frac{1}{2}$ subsets are appended together in a certain way, then the resultant subset must be strictly rate $\frac{1}{2}$. This can also be seen as a general necessary condition for rate $\frac{1}{3}$ feasibility based on the properties of strictly rate $\frac{1}{2}$ sets and their intersections.

Theorem 3. Let $\mathcal{W}_i : i = 1, 2, \dots, r$ ($r \geq 2$) be strictly rate $\frac{1}{2}$ sets of a rate $\frac{1}{3}$ feasible index coding problem \mathbb{I} with message set \mathcal{W} , such that the following holds.

- The sets $(\cup_{i=1}^{t-1} \mathcal{W}_i) \cap \mathcal{W}_t, 2 \leq t \leq r$, are strictly rate 1 infeasible.

Then the set $\cup_{i=1}^r \mathcal{W}_i$ must be strictly rate $\frac{1}{2}$.

Thus, if the index coding problem \mathbb{I} restricted to $(\cup_{i=1}^r \mathcal{W}_i)$ has any $(\cup_{i=1}^r \mathcal{W}_i)$ -restricted internal conflicts, then \mathbb{I} is not rate $\frac{1}{3}$ feasible.

Proof: If we are given that $\cup_{i=1}^r \mathcal{W}_i$ is strictly rate $\frac{1}{2}$, then the claim about rate $\frac{1}{3}$ infeasibility follows from the necessary condition for rate $\frac{1}{2}$ feasibility in [8]. Hence, it is enough to show that $\cup_{i=1}^r \mathcal{W}_i$ is strictly rate $\frac{1}{2}$. Let \mathbb{E} be the encoding function of some $\frac{1}{3}$ index code. It is clear that $\dim(V_{\mathbb{E}}(\cup_{i=1}^r \mathcal{W}_i)) \geq 2$, as each set \mathcal{W}_i is strictly rate $\frac{1}{2}$. We prove the claim by induction on r . The claim is true for $r = 1$ by assumption that each set \mathcal{W}_i is strictly rate $\frac{1}{2}$. Now

suppose that the statement is true upto $r - 1$. Then we have $\cup_{i=1}^{r-1} \mathcal{W}_i$ and \mathcal{W}_r are strictly rate $\frac{1}{2}$ and $(\cup_{i=1}^{r-1} \mathcal{W}_i) \cap \mathcal{W}_r$ is strictly rate 1 infeasible. This means that $(\cup_{i=1}^{r-1} \mathcal{W}_i) \cap \mathcal{W}_r$ must be strictly rate $\frac{1}{2}$ feasible as \mathcal{W}_r is strictly rate $\frac{1}{2}$. Thus, we have $V_{\mathbb{E}}(\cup_{i=1}^{r-1} \mathcal{W}_i) = V_{\mathbb{E}}(\mathcal{W}_r) = V_{\mathbb{E}}((\cup_{i=1}^{r-1} \mathcal{W}_i) \cap \mathcal{W}_r)$. Thus $V_{\mathbb{E}}((\cup_{i=1}^{r-1} \mathcal{W}_i) \cup \mathcal{W}_r) = V_{\mathbb{E}}(\cup_{i=1}^{r-1} \mathcal{W}_i)$ as well, and hence $(\cup_{i=1}^{r-1} \mathcal{W}_i) \cup \mathcal{W}_r$ must be strictly rate $\frac{1}{2}$. ■

Now, we will give three examples of sets (Corollary 1, 2, and 3) which satisfy the condition in Theorem 3.

Corollary 1 (Result from [11]). *Any type-2 alignment set of an index coding problem \mathbb{I} is strictly rate $\frac{1}{2}$.*

Proof: Let $\mathcal{W}_i : i = 1, \dots, r$ be the set of all triangular interfering sets in the given type-2 set. Clearly each one of them is strictly rate $\frac{1}{2}$. Furthermore, as the type-2 set is a maximally connected set of triangular interfering sets, we can assume WLOG that the sets \mathcal{W}_i are in an order such that $(\cup_{i=1}^{t-1} \mathcal{W}_i) \cap \mathcal{W}_t, 2 \leq t \leq r$ contains a conflict, and thus is strictly rate 1 infeasible. The corollary follows by applying Theorem 3. ■

Definition 11 (Extended Type-2 Alignment Set). *Consider $\mathcal{W}_i : i = 1, \dots, r$ be a maximal collection of type-2 alignment sets such that $(\cup_{i=1}^{t-1} \mathcal{W}_i) \cap \mathcal{W}_t, 2 \leq t \leq r$ contains at least one conflict. Then the union $\cup_{i=1}^r \mathcal{W}_i$ is termed as extended type-2 alignment set (Xtype-2 set in short).*

We then have the following corollary, which we state without proof as it is a direct application of Theorem 3.

Corollary 2. *Any Xtype-2 set of an index coding problem \mathbb{I} is strictly rate $\frac{1}{2}$.*

For our final corollary, we define SPIC alignment sets as follows, similar to type-2 alignment sets (Definition 7).

Definition 12 (Adjacent SPIC sets, SPIC Alignment Set). *Two SPIC subsets of \mathcal{W}_1 and \mathcal{W}_2 are said to be adjacent SPIC subsets if $\mathcal{W}_1 \cap \mathcal{W}_2$ is strictly rate 1 infeasible. The SPIC sets \mathcal{W}_1 and \mathcal{W}_2 are called connected if there exists a sequence of adjacent SPIC sets from \mathcal{W}_1 to \mathcal{W}_2 . A SPIC alignment set is a maximal set of SPIC subsets which are connected to each other.*

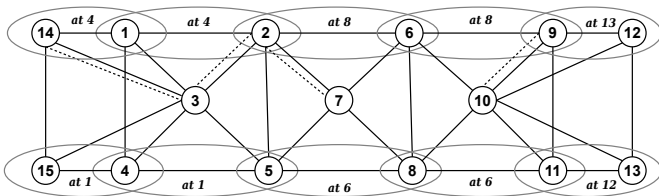


Fig. 3. A SPIC Alignment set. The sets $\{14, 1, 3, 15, 4\}$, $\{2, 1, 3, 5, 4\}$, $\{2, 6, 7, 5, 8\}$, $\{9, 6, 10, 11, 8\}$ and $\{9, 12, 10, 11, 13\}$ are all SPIC sets. The pair-wise intersections in the same sequence are $\{1, 3, 4\}$, $\{2, 5\}$, $\{6, 8\}$ and $\{9, 10, 11\}$ respectively, and all such intersections are strictly rate 1 infeasible. The messages within the ellipses interfere at the messages as mentioned in the ellipse.

Fig. 3 shows an example shows an example of a SPIC alignment set. We now give the related result.

Corollary 3. *A SPIC alignment set of \mathbb{I} is strictly rate $\frac{1}{2}$.*

Proof: Let \mathcal{W}_a be the SPIC alignment set. Let $\mathcal{W}_i : i = 1, \dots, r$ be the set of all SPIC sets which comprise \mathcal{W}_a , ordered such that \mathcal{W}_t is connected to $(\cup_{i=1}^{t-1} \mathcal{W}_i), 2 \leq t \leq r$ (we can assume this WLOG as \mathcal{W}_a is maximally connected set of SPIC sets). The corollary follows by applying Theorem 3. ■

The sets described in Corollaries 1, 2, 3 are only a few examples of the possible configurations of strictly rate $\frac{1}{2}$ sets in \mathbb{I} . We can imagine several other configurations as well, involving a mixture of Xtype-2 sets, SPIC sets, and other strictly rate $\frac{1}{2}$ sets (if they exist), satisfying the conditions in Theorem 3, leading to bigger structures which must be strictly rate $\frac{1}{2}$. An interesting open question is to determine if there a finite number of basic substructures, like the triangular interfering sets, SPIC sets, such that their maximally connected versions should satisfy specific properties for rate $\frac{1}{3}$ feasibility.

V. A SUFFICIENT CONDITION FOR RATE $\frac{1}{3}$ FEASIBILITY

In this section, we introduce the notion of contraction of an alignment edge and contraction of an index coding problem. We prove that an index code for the contraction of an index coding problem can be extended to that of the original index coding problem. We also give a sufficient condition in terms of Xtype-2 sets in the maximal contraction of an index coding problem for the original index coding problem to be rate $\frac{1}{3}$ feasible.

Definition 13 (Contraction of an alignment edge). *An alignment edge connecting two vertices (messages) W_i and W_j which are not in conflict is said to be contracted by identifying the vertices as a single vertex thereby making all the edges (alignment edges and conflict hyperedges) that were incident on W_i and W_j now incident on the newly created vertex.*

When we contract any alignment edge, note that we get a derived index coding problem with one less the number of messages. The derived index coding problem is completely characterised by the contracted alignment graph and the correspondingly altered conflict graph. We can envision a sequence of such derived index coding problems, obtained by a sequence of contractions of the alignment set of the original problem, each with one lesser message than the previous. Abusing the definition, an index coding problem \mathbb{I}' is said to be a contraction of an index coding problem \mathbb{I} if it is obtained by a finite sequence of contractions of the alignment edges of \mathbb{I} . The following definition captures the maximal cases after which we cannot contract anymore.

Definition 14 (Maximal Contraction of an Index Coding Problem). *Let \mathbb{I} be a given index coding problem. Let \mathbb{I}_{max} be an index coding problem obtained by a sequence of contractions of \mathbb{I} , such that any two messages of \mathbb{I}_{max} connected by an alignment edge are also in conflict in \mathbb{I}_{max} . Then \mathbb{I}_{max} is said to be a maximal contraction of \mathbb{I} .*

Clearly there could be multiple distinct maximal contractions of a given index coding problem, depending on the sequence of alignment edges which are contracted.

Lemma 3. *Let \mathbb{I}' be a contraction of \mathbb{I} . Any scalar linear index coding solution of rate R for \mathbb{I}' can be extended to a scalar linear index coding solution of rate R for \mathbb{I} .*

Proof: WLOG, we assume that \mathbb{I}' is obtained from \mathbb{I} by a single contraction of an alignment edge. Let $\mathbf{W} = (W_1, \dots, W_n)$ denote the vector of messages in \mathbb{I} .

WLOG, consider that symbols W_{n-1} and W_n in \mathbb{I} were contracted to get \mathbb{I}' . We denote the new symbol (vertex) which is created as the message W'_{n-1} . The vector of messages, the demand sets, the side-information sets and the interfering sets of \mathbb{I}' (denoted as $\mathbf{W}', D'(j), S'(j)$, and $Interf'_j(k)$ respectively) can be obtained from those of \mathbb{I} by replacing $W_i : 1 \leq i \leq n-1$ with $W'_i : 1 \leq i \leq n-1$, and replacing W_n with W'_{n-1} . We note that the number of receivers remains unchanged in \mathbb{I}' .

Let \mathbb{E}' be the encoding function of a solution to \mathbb{I}' . Thus we must have

$$V'_k \notin V'_{\mathbb{E}'}(Interf'_j(k)), \quad (1)$$

for all $W'_k \in D(j)$, \forall receivers j , where V'_k is the vector which is assigned to W'_k in \mathbb{E}' .

Consider the encoding function \mathbb{E} for \mathbb{I} defined as follows.

$$V_k = \begin{cases} V'_k & \forall 1 \leq k \leq n-1. \\ V'_{n-1} & \text{if } k = n. \end{cases}$$

We will now show that this assignment resolves all conflicts. Consider a message W_k for some $k \leq n-2$ demanded at some receiver j . Thus $V_k = V'_k$. For some message W_k demanded at receiver j , suppose $Interf_j(k)$ does not contain W_n . Then $Interf_j(k) = Interf'_j(k)$ (except for the ‘dashes’). If $Interf_j(k)$ contains W_n , then $Interf'_j(k)$ contains W'_{n-1} in the place of W_n , apart from having other ‘dashed’ messages corresponding to those in $Interf_j(k)$. In any case, we must have

$$V_{\mathbb{E}}(Interf_j(k)) = V_{\mathbb{E}'}(Interf'_j(k)), \forall j, \forall k \leq n-2. \quad (2)$$

By (1) and (2), we have $V_k \notin V_{\mathbb{E}}(Interf_j(k))$.

Now we consider W_{n-1} (equivalently, W_n) demanded at receiver j . The receiver j in \mathbb{I}' demands the symbol W'_{n-1} , and it is given that (1) holds. Clearly, $Interf'_j(n-1)$ does not contain W'_{n-1} . This means that $Interf'_j(n-1)$ (equivalently, $Interf_j(n)$) does not contain W_n or W_{n-1} . Thus, by similar arguments as above, we must have

$$V_{\mathbb{E}}(Interf_j(n-1)) = V_{\mathbb{E}'}(Interf'_j(n-1))$$

(equivalently, $V_{\mathbb{E}}(Interf_j(n)) = V_{\mathbb{E}'}(Interf'_j(n-1))$). By assignment \mathbb{E} , we have $V_{n-1} \notin V_{\mathbb{E}}(Interf_j(n-1))$ (equivalently, $V_n \notin V_{\mathbb{E}}(Interf_j(n))$). Thus all the conflicts in \mathbb{I} are resolved. This concludes the proof. ■

Theorem 4. *A rate $\frac{1}{2}$ infeasible index coding problem \mathbb{I} is rate $\frac{1}{3}$ feasible if there exists a maximal contraction \mathbb{I}' of the index coding problem \mathbb{I} such that the following conditions hold:*

- (a) *Any Xtype-2 set in \mathbb{I}' has no restricted internal conflicts.*
- (b) *No three Xtype-2 sets have a message vertex in common.*

- (c) *For any distinct Xtype-2 sets $\mathcal{W}_i, \mathcal{W}_j$, if $\mathcal{W}_i \cap \mathcal{W}_j \neq \phi$, then there is no conflict between any two messages in $\mathcal{W}_i \cap \mathcal{W}_j$.*

Proof: To prove the theorem, we will give a solution to index coding problem \mathbb{I}' and then apply Lemma 3 to extend the solution to the actual index coding problem \mathbb{I} . For notational convenience, we will drop the dash associated with the variables corresponding to the index coding problem \mathbb{I}' (We will however refer to index coding problem as \mathbb{I}'). Let $\mathcal{W}_i, 1 \leq i \leq r$ denote the Xtype-2 sets of \mathbb{I}' . By assumption, we have $\mathcal{W}_i \cap \mathcal{W}_j \cap \mathcal{W}_k = \phi$, for all distinct i, j, k . Note that the condition (c) implies that there is no alignment edge between any two messages in $\mathcal{W}_i \cap \mathcal{W}_j$, as \mathbb{I}' is a maximal contraction.

Consider a graph which has r vertices, where each vertex represents an Xtype-2 set. There is an edge between two vertices in this graph, if the two Xtype-2 sets intersect in one or more messages. We will refer to this graph as the Extended Type-2 Intersection Graph (ETIG). We will first assign vectors to the edges in ETIG and then use that in turn to come up with an assignment for the messages in the index coding problem \mathbb{I}' .

The algorithm to assign vectors to the edges of the ETIG is as follows. We start with the assumption that the edge set is non-empty, else the algorithm terminates straightaway. We repeat the below steps until all edges have been assigned vectors.

- 1: Pick an unassigned edge e_{ij} between vertices i and j . Let \mathcal{V}_i denote the set of vectors already assigned to edges incident on vertex i and similarly \mathcal{V}_j .
- 2: Suppose both $sp(\mathcal{V}_i)$ and $sp(\mathcal{V}_j)$ are not 2-dimensional, then assign a random 3×1 vector to the edge.
- 3: Suppose exactly one of $sp(\mathcal{V}_i)$ or $sp(\mathcal{V}_j)$ is 2-dimensional. For example, let $sp(\mathcal{V}_i)$ be 2-dimensional. Then, a random 3×1 vector from $sp(\mathcal{V}_i)$ is assigned to the edge.
- 4: Suppose both $sp(\mathcal{V}_i)$ and $sp(\mathcal{V}_j)$ are 2-dimensional. Then, a 3×1 vector from the intersection $sp(\mathcal{V}_i) \cap sp(\mathcal{V}_j)$ is assigned to the edge. Such a vector always exists in the intersection since both $sp(\mathcal{V}_i)$ and $sp(\mathcal{V}_j)$ are 2-dimensional subspaces of a 3-dimensional space.

We note that at the end of the algorithm, each vertex with degree at least one in ETIG has a set of vectors associated with edges incident on them, the span of which is either 1-dimensional or 2-dimensional. An example ETIG and assignment of vectors to the edges in the graph is illustrated in Fig. 4.

We now describe the procedure to assign vectors to the messages in the index coding problem \mathbb{I}' .

- 1) Firstly, vectors are assigned to the messages in the intersections of Xtype-2 sets. We assign the same vector to all the messages in the intersection $\mathcal{W}_i \cap \mathcal{W}_j$. We pick the vector which was assigned to the edge which joined the vertices corresponding to i^{th} and j^{th} Xtype-2 set in ETIG and assign it to all the messages in the intersection $\mathcal{W}_i \cap \mathcal{W}_j$. This step is repeated for all intersections of Xtype-2 sets.

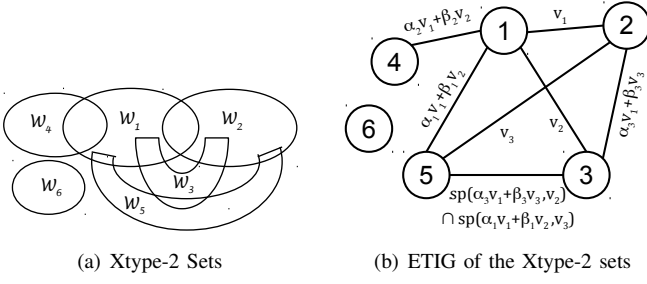


Fig. 4. The example has 6 Xtype-2 sets with the intersections as shown. The ETIG of the Xtype-2 sets is given. To assign vectors, we start by assigning v_1 to edge e_{12} and v_2 to edge e_{13} . Then edges e_{14} and e_{15} are assigned random linear combinations of v_1 and v_2 . Next, e_{25} is assigned v_3 . e_{23} is assigned random linear combination of v_1 and v_3 . Finally, since both vertices 3 and 5 are already assigned 2-D spaces, a vector in the intersection of the two spaces is assigned to e_{35} .

2) At the end of the above step, the messages in \mathcal{W}_i s maybe partially assigned. To assign vectors to remaining messages (if such non-assigned messages exist) in the Xtype-2 sets \mathcal{W}_i s, we consider the following three subcases:

- For any i , if there is just one index j such that $\mathcal{W}_i \cap \mathcal{W}_j \neq \emptyset$, then we have assigned only one vector to \mathcal{W}_i at the end of step 1). Let the vector be denoted by $\underline{v}_1^{(i)}$. Now, we pick another 3×1 random vector $\underline{v}_2^{(i)}$. Then, we assign a random vector in the span of $\underline{v}_1^{(i)}$ and $\underline{v}_2^{(i)}$ to each message in \mathcal{W}_i .
- For any i , if there is more than one index j such that $\mathcal{W}_i \cap \mathcal{W}_j \neq \emptyset$, then we have assigned a 2-dimensional space to messages in \mathcal{W}_i at the end of step 1). To each of the remaining messages in \mathcal{W}_i , we assign a random vector from the same 2-dimensional space.
- For any i , if there is no index j such that $\mathcal{W}_i \cap \mathcal{W}_j \neq \emptyset$ (corresponds to isolated vertices in ETIG), then we first pick a randomly generated 2-dimensional space. Then, we assign a random 3×1 vector in the 2-D space to each message in the i^{th} Xtype-2 set.

3) For all other messages in \mathbb{I}' (not in Xtype-2 sets), we assign a random 3×1 vector.

Let \mathbb{E} denote the encoding function corresponding to this assignment and V_k denote the vector assigned to the message W_k . We now show that this assignment resolves all conflicts in \mathbb{I}' , i.e., with high probability $V_k \notin V_{\mathbb{E}}(\text{Interf}_k(j))$. Consider a receiver j which requests a message W_k . We will classify all the conflicts into the following cases and verify that conflicts are resolved in each case.

- If $|\text{Interf}_k(j)| = 1$, then the conflict is resolved, since any two messages which are in conflict are assigned linearly independent vectors according to the scheme. It is only the intersections of Xtype-2 sets which are assigned the same vector and we have already noted earlier that all the messages in any given intersection do not have any conflicts within them.
- Consider the case when $|\text{Interf}_k(j)| = 2$. The two messages within the set $\text{Interf}_k(j)$ should have a conflict between them since \mathbb{I}' is the maximal contraction. Now, we have the following two sub cases:

- Both W_k and $\text{Interf}_k(j)$ belong to a Xtype-2 alignment set. This is not possible since then the Xtype-2 set will have a restricted internal conflict.
- For any configuration of W_k and $\text{Interf}_k(j)$ other than the case considered above, the conflict is resolved. This is true because any three messages which do not all belong to an Xtype-2 alignment set and which have conflicts between them are assigned three linearly independent vectors according to the scheme. Here again, note that $\text{Interf}_k(j)$ cannot be contained within the intersection of two Xtype-2 sets (since all the messages in any given intersection do not have any conflicts within them).
- Consider the case when $|\text{Interf}_k(j)| \geq 3$. We note that in the maximal contraction \mathbb{I}' , any triangular interfering set of \mathbb{I}' has to be contained within a type-2 alignment set and hence within a Xtype-2 set. Thus, $\text{Interf}_k(j)$ is contained within a Xtype-2 set. Also since Xtype-2 sets of \mathbb{I}' are assumed to have no restricted internal conflicts, W_k has to be outside the Xtype-2 set. Since $\text{Interf}_k(j)$ lies within a Xtype-2 set and the messages within a Xtype-2 set are assigned vectors from a two dimensional space, $V_{\mathbb{E}}(\text{Interf}_k(j))$ is a two dimensional space. Noting that V_k is a random 3×1 vector, we have that $V_k \notin V_{\mathbb{E}}(\text{Interf}_k(j))$ with high probability. Hence, the conflict is resolved.

Thus, all the conflicts in \mathbb{I}' are resolved. Now, the solution given above can be extended to give a solution of the original index coding problem \mathbb{I} by 'uncontracting' one alignment edge at a time and doing an assignment as described in the proof of Lemma 3. This concludes the proof. ■

Remark 2. We claim that Theorem 2 can be seen as a special application of Theorem 4 (which is the most general result known so far for rate $\frac{1}{3}$ feasible problems), and this can be shown using the proof details of Theorem 2 which is available in [11]. We leave the details to the reader.

REFERENCES

- Y. Birk and T. Kol, "Coding on Demand by an Informed Source (ISCOD) for Efficient Broadcast of Different Supplemental Data to Caching Clients", IEEE Transactions on Information Theory, Vol. 52, No. 6, June, 2006, pp. 2825-2830.
- Z.Bar-Yossef, Y. Birk, T.S. Jayram, T. Kol, "Index Coding with Side Information", IEEE Transactions on Information Theory, Vol. 57, No. 3, March 2011, pp. 1479-1494.
- L. Ong, C. K. Ho, and F. Lim, "The single-uniprior index-coding problem: The single-sender case and the multi-sender extension," IEEE Transactions on Information Theory, Vol. 62, No. 6, June 2016, pp. 3165-3182.
- K. Shanmugam, A. G. Dimakis, and M. Langberg, "Local graph coloring and index coding", Proceedings of IEEE International Symposium on Information Theory (ISIT) 2013, Istanbul, Turkey, Jul 7-12, pp. 1152-1156.
- A. Blasiak, R. Kleinberg, and E. Lubetzky, "Index Coding via Linear Programming", ArXiv-CoRR, Jul. 2011, Available at <http://arxiv.org/abs/1004.1379>.
- A. Blasiak, R. Kleinberg, and E. Lubetzky, "Broadcasting with side information: Bounding and approximating the broadcast rate", IEEE Transactions on Information Theory, Vol. 59, No. 9, Sep. 2013, pp. 5811-5823.
- C. Thapa, L. Ong, S. J. Johnson, "Generalized interlinked cycle cover for index coding", IEEE ITW (Fall) 2015, Jeju Island, South Korea, Oct. 11-15 2015.

- [8] S. A. Jafar, "Topological Interference Management Through Index Coding", IEEE Transactions on Information Theory, Vol. 60, No. 1, Jan. 2014, pp. 529-568.
- [9] H. Maleki, V. R. Cadambe, and S. A. Jafar, "Index Coding-An Interference Alignment Perspective", IEEE Transactions on Information Theory, Vol. 60, No. 9, Sep. 2014, pp. 5402-5432.
- [10] R. Peeters, "Orthogonal representations over finite fields and the chromatic number of graphs", Combinatorica, Vol. 16, No. 3, pp. 417-431, Sept 1996.
- [11] P. Krishnan and V. Lalitha, "A class of index coding problems with rate $\frac{1}{3}$ ", Proceedings of IEEE International Symp. on Info. Theory (ISIT), 2016, July 10-15, Barcelona, Spain, pp. 130-134.